MINIMUM INSERTION LOSS FILTERS

by

E. G. Fubini*

and

E. A. Guillemin**

INTRODUCTION

In microwave applications, it is important to minimize the midband insertion loss of filters. This requirement represents one of the most important differences between the design of low-frequency filters and microwave filters and is due to the fact that low-noise-figure amplification is a problem at microwave frequencies.

Since microwave filters have been in existence for a short time, the problem of losses played a secondary role in the theory of filter synthesis; only recently has the problem even been considered.

The purpose of this paper is to:

a. Assign the minimum midband loss realizable in a band-pass (or low-pass) filter of the Butterworth or Tchebycheff type,

b. Show that two universal curves give satisfactory approximations for all bandwidths, quality of components, number of sections, and acceptable ripple in the pass band,

c. Show that a filter with minimum loss can actually be designed, and

d. Prove that Tchebycheff filters are very undesirable from the point of view of losses.

USE OF DESIGN CURVES

This paper refers to the case of band-pass filters, including the low-pass prototypes, whose bandwidth is sufficiently narrow for the coupling between sections to be considered independent of frequency. In this case, Dishal has shown that exact Butterworth (flat-flat) and Tchebycheff shapes can be obtained, provided that the loss in each section (considered isolated and unloaded) does not exceed a specific amount. It is convenient to use the unloaded $Q = Q_{unl}$ of each section as a measure of this loss. Dishal has given the values that the unloaded Q must exceed for a given set of specifications to be realizable.

Figure 1 gives this minimum unloaded $Q = Q_{min}$. Note that the ordinates are normalized to unity bandwidth--that is, instead of plotting Q_{min}, a more general set of curves has been obtained by plotting

$$q_{min} = Q_{min} \frac{\Delta f}{f_o} = \frac{1}{\text{figure of merit}}$$

where $\Delta f/f_o$ is the relative bandwidth of the complete filter connected to the proper load and generator. In Figure 1, note how, for a given bandwidth, the requirements in component quality (higher q_{min}) increase sharply from the no-ripple (Butterworth) case to the cases where ripples of a few decibels are specified. It must also be remembered that, if one were trying to use components whose unloaded Q (Q_{unl}) is barely equal to the required minimum,

$$Q_{min} = \frac{f_o}{\Delta f} q_{min}$$

the desired filter shape could indeed be achieved <u>but with an infinite midband loss</u>.

The question then arises:

<u>If the unloaded Q of each filter section is</u>

$$Q_{unl} > Q_{min}$$

<u>what is the minimum midband loss L?</u> Up to this point, all we have stated is that using the minimum acceptable Q produces infinite losses; however, one question remains unanswered: what quality can we expect in a filter as a function of the losses in the individual sections? The answer for the case of three sections was given by Taub and Bogner[1] in graphical form and by Fubini[1] in analytical form.

* Airborne Instruments Laboratory Mineola, New York.

** Massachusetts Institute of Technology Cambridge, Mass.

1. J. J. Taub and B. F. Bogner, Proc. IRE, Vol. 45, p 681-687, May 1957.

The solution for the case of four sections was calculated by Aylward[2] using a digital computer. Here, we want to show the surprising fact that the answer does not primarily depend upon the number of sections and is almost exclusively controlled by the value of the ratio

$$x = \frac{Q_{unl}}{Q_{min}}$$

between the unloaded Q of the components used and the minimum acceptable value of such Q. In other words, once the minimum value of Q_{unl} is obtained from Figure 1, and the quality of available components is determined by their unloaded Q, the loss L is almost completely defined and varies very little with the number of sections, the shape of the filter, the bandwidth, etc. It must be remembered that this statement is not exactly true, and it is definitely incorrect if shapes other than Butterworth or Tchebycheff are considered.

Figure 2 shows a set of curves that give the minimum insertion loss L (db) of Butterworth filters plotted as a function of the ratio x previously defined. The curves are labeled with a number, which indicates the number of sections. It is clear that, at least for moderate losses, the curves are very close to each other. The curves for one and two sections are exactly the same and given by the equation

$$L = 20 \log_{10} \frac{x}{x - 1}$$

Although the curves for Tchebycheff filters are not shown, they fall in close proximity to the corresponding Butterworth curves.

Figures 3 and 4 show a group of curves of loss (db) plotted vs

$$Q_{unl} \frac{\Delta f}{f_o} = Q_u \frac{\Delta f}{f_o} \text{ and } x = \frac{Q_{unl}}{Q_{min}}$$

The surprising relation between such curves, already shown in Figure 1, is shown again in Figure 4. It would be desirable if a single curve could be substituted for the family of curves shown in Figures 2 and 4. This could be accomplished with fair approximation using the two curves shown in Figures 5 and 6. The first curve (Figure 5) is the same as the curve plotted in Figure 2 for one and two sections (see equation 1). Figure 6 introduces a correction in Figure 5 by introducing a multiplying factor that depends upon the number of sections. Although the data shown in Figure 6 are not exact, they greatly improve the accuracy of the data shown in Figure 5; however, this is only approximately true. These two curves (Figures 5 and 6) are completely sufficient for all practical purposes. For example, let a filter be required with a flat-flat response and six sections. (Note that this filter must have a minimum midband loss and a relative bandwidth of 1 percent.) Assume that the unloaded Q of each section is 2000. From Figure 1 we find that, for a Butterworth type of six-section filter

$$q_{min} = Q_{min} \frac{\Delta f}{f} = \frac{1}{100} Q_{min} = 3.86$$

The value of x can now be computed, as follows:

$$x = \frac{Q_{unl}}{Q_{min}} = \frac{2000}{386} = 5.18$$

Entering Figure 5 with this value of x, one finds L = 1.91 db, which is the minimum insertion loss for one and two sections. In our case, since the required number of sections is six, we find (from Figure 6) that the loss is higher by a correction factor of 1.23. We conclude that the filter will have a minimum midband loss of

$$L = 1.91 \times 1.23 = 2.35 \text{ db}$$

which differs only by 5 percent* or 0.1 db from the value L = 2.25 db given by accurate computations and noted in Figure 2.

OPTIMUM FILTERS

Figure 5 shows that, for a given quality of components, filters with the lowest q_{min} have the minimum midband loss. Figure 6 shows that, for filters with six sections or less, the correction due to the number of sections does not exceed 25 percent. For these reasons, $1/q_{min}$ can be considered as a figure of merit; it is interesting to compare several filters to determine those that are best from the point of view of minimum loss.

This has been accomplished in Figure 7, where the figure of merit of a group of filters has been plotted as a

2. W. R. Aylward, unpublished thesis.

* This is a typical error. The use of Figure 6 does not introduce errors greater than 7 percent in db for a filter with less than 20 sections.

function of the skirt selectivity. Each curve corresponds to a given number of sections, and each point of the curves corresponds to Tchebycheff filters of variable ripple. The upper points of every curve correspond to Butterworth filters. It is clear that, for a given skirt selectivity, the figure of merit is higher and therefore the losses lower if selectivity is obtained by increasing the number of sections and not increasing the ripples. In this respect, Tchebycheff filters are much worse than Butterworth filters. A curve can be imagined that joins all the upper points in the curves shown in Figure 7. The curve obtained by means of this procedure supplies information as to the best possible figure of merit for this type of filter for a given skirt selectivity.

THEORETICAL CONSIDERATIONS

It is well known that the effect of losses can be eliminated from the synthesis process in the case of identical sections: a Butterworth filter, for example, can be obtained by requiring that the poles lie on a circle whose center is not on the frequency axis but whose distance from this axis is equal to the unloaded decrement of the elements of the filter. As shown in Figure 8, if the filter were lossless, ℓ would be the frequency axis and the response would be flat-flat. Because of the losses, the poles had to be predistorted; this operation is equivalent to that of taking f as the frequency axis. Obviously, the response along f will be peaked, and the maximum value of this peak cannot exceed unity. This is equivalent to the statement that, in the response

$$t(s) = \frac{K}{(s - s_1)(s - s_2) \ldots (s - s_n)}$$

K must be smaller than the maximum value K_o of $\Pi(s - s_i)$. Therefore, this maximum value K_o represents the maximum filter efficiency and $1/K_o$ represents the minimum loss.

Let the Butterworth polynomial be

$$P(s) = (s - s_1)(s - s_2) \ldots (s - s_n) \qquad (1)$$

with

$$s_v = j\, e^{j(2v - 1)\pi/2n} \qquad (2)$$

where v = 1, 2, ... n.

Write

$$P^*(s) = (s - s_2)(s - s_3) \ldots (s - s_n) \qquad (3)$$

so that

$$P(s) = (s - s_1)P^*(s) \qquad (4)$$

The horizontal distance from s_1 to the j-axis is

$$a = \sin \pi/2n \qquad (5)$$

At the point $(s_1 + a)$ on the j-axis, we assume that the magnitude of P(s) is very nearly unity, because this point is still in the pass band. Equation 4 therefore gives

$$|P(s_1 + a)| \approx 1 \approx a\,|P^*(s_1 + a)|$$

or

$$|P^*(s_1 + a)| \approx 1/a = = 1/\sin \pi/2n \approx 2n/\pi \text{ for large } n \qquad (6)$$

If we assume that P*(s) does not vary appreciably in the interval between $s = s_1$ and $s = s_1 + a$ (to be justified later), then in equation 4 we can replace $|P^*(s)|$ by the constant value (given in equation 6) and have

$$\frac{1}{|P(s)|} = \frac{a}{|s - s_1|} \qquad (7)$$

which yields unity (as it should) at $s = s_1 + a$. If we write $s = s_1 + a - \delta$, where δ is the shift of the pole pattern, then within the interval $s = s_1$ to $s = s_1 + a$ we have (for the Butterworth function) an approximate representation that reads

$$\frac{1}{|P(s)|} = \frac{a}{a - \delta} = \frac{1}{1 - (\delta/a)} \qquad (8)$$

Observe that this relation gives the correct value (unity) for $\delta = 0$ regardless of the order n, and approximately correct values for $0 \leq \delta < a$ to the extent that P*(s) in equation 3 is constant within this interval. To investigate the latter, let us compute the ratio $P^*(s_1)/P^*(s_1 + a)$, or since $s_1 + a = j$, the ratio $P^*(s_1)/P^*(j)$.

From equation 3, we find in a straightforward manner

$$\left(\frac{dP^*}{ds}\right)_{s=j} = P^*(j) \sum_{v=2}^{n} \frac{1}{j - s_v} = \frac{2n}{\pi} \sum_{v=2}^{n} \frac{1}{j - s_v} \tag{9}$$

where equation 6 is made use of in the last step. If we use the relation

$$P^*(s_1) = P^*(j) - a\left(\frac{dP^*}{ds}\right)_{s=j} \tag{10}$$

and substitute for $\underline{a}$ the value $\pi/2n$, we have

$$\frac{P^*(s_1)}{P^*(j)} = 1 - \left(\frac{\pi}{2n}\right)^2 \left(\frac{dP^*}{ds}\right)_{s=j} = 1 - \frac{\pi}{2n} \sum_{v=2}^{n} \frac{1}{j - s_v} \tag{11}$$

Substituting the value of s_v (equation 2),

$$x = (2v - 1)\ \pi/2n, \tag{12}$$

and noting that

$$\frac{1}{j(1 - e^{jx})} = \frac{1}{2j}\left(1 + \cot\frac{x}{2}\right) = \frac{1}{2}\left(\cot\frac{x}{2} - j\right), \tag{13}$$

equation 11 becomes

$$\frac{P^*(s_1)}{P^*(j)} = 1 - \frac{\pi}{4n} \sum_{v=2}^{n} \left(\cot\frac{x}{2} - j\right) = 1 + j\,\frac{(n-1)\pi}{4} - \frac{1}{2} \sum_{\frac{x}{2}=\frac{3\pi}{4n}}^{\frac{x}{2}=\frac{\pi}{2}-\frac{\pi}{4n}} \cot\frac{x}{2}\,\frac{\Delta x}{2} \tag{14}$$

where $\Delta x = \pi/n$. From a graphical sketch of the contangent function and its step approximation in equation 14, we see that a very close value for this sum may be obtained from the integral:

$$\frac{1}{2}\int_{\pi/2n}^{\pi/2} \cot X\, dx = \frac{1}{2}\,\ell n\,(\sin x)\Big]_{\pi/2n}^{\pi/2} = \frac{1}{2}\,\ell n\,\frac{2n}{\pi} \tag{15}$$

and so we finally obtain

$$\frac{P^*(s_1)}{P^*(j)} = \left(1 - \frac{1}{2}\,\ell n\,\frac{2n}{\pi}\right) + j\,\frac{(n-1)\pi}{4n} \tag{16}$$

The magnitude of this ratio for various values of n is tabulated below:

n =	3	6	10	15
ratio =	0.855	0.733	0.708	0.744

n =	20	30	40	50
ratio =	0.793	0.898	0.984	1.06

These results show that the assumption made in the derivation of the approximate result given in equation 8 is entirely reasonable for any values of n to be encountered in a practical problem. One should observe also that the result (given in equation 8) is exact for $\delta = 0$ and has its greatest degree of approximation near $\delta = a$, where the Butterworth function blows up. Such values for the shift δ will hardly be used in a practical problem.

The design of the filter can be now described as follows.

Starting from the normal unshifted Butterworth transfer function $t(s) = 1/P(s)$, where $P(s)$ is the usual Butterworth polynomial, we form the shifted and K-multiplied function $Kt(s - \delta)$, where K is determined so that the absolute value of this function for $s = j$ never exceeds unity. This is the condition that fixes the largest possible value of K (least midband loss) that can yield a positive real $Z_1(s)$-function. Our approximate formula determines this largest K value; and the necessary positive-real character of $Z_1(s)$ is the reason that the resulting midband loss is minimum. $Z_1(s)$ is determined from the relation

$$Z_1(s) = \frac{1 - \rho(s)}{1 + \rho(s)}$$

where $\rho(s)$ is obtained from

$$|\rho(j\omega)|^2 = 1 - K|t(j\omega - \delta)|^2$$

The $Z_1(s)$ function thus obtained is the desired preshifted impedance function. It is realized in the normal manner by a lossless network terminated in resistance. The incidental loss that is present in the actual network then shifts the pole pattern of $Kt(s)$ back to the original Butterworth position.

SUMMARY

A generalized criterion is given for the design of minimum loss Butterworth and Tchebycheff filters of arbitrary bandwidth and component quality. This criterion is chosen to minimize the insertion loss in the center of the band and is particularly useful for microwave filters, where insertion loss must be minimized. Two general curves are given for this design, together with information as to the method followed in obtaining them. It is shown that, if insertion losses are to be minimized, the use of Tchebycheff filters with ripples greater than very small fractions of one decibel must be discouraged.

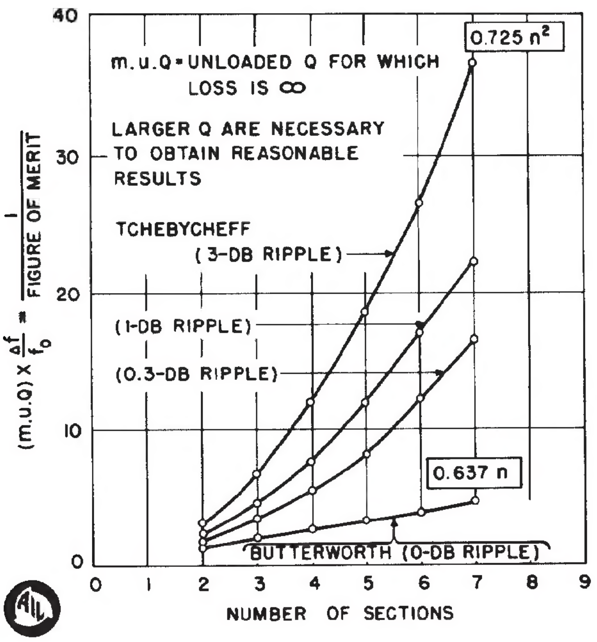

Fig. 1 Minimum unloaded Q for Butterworth and Tchebycheff filters.

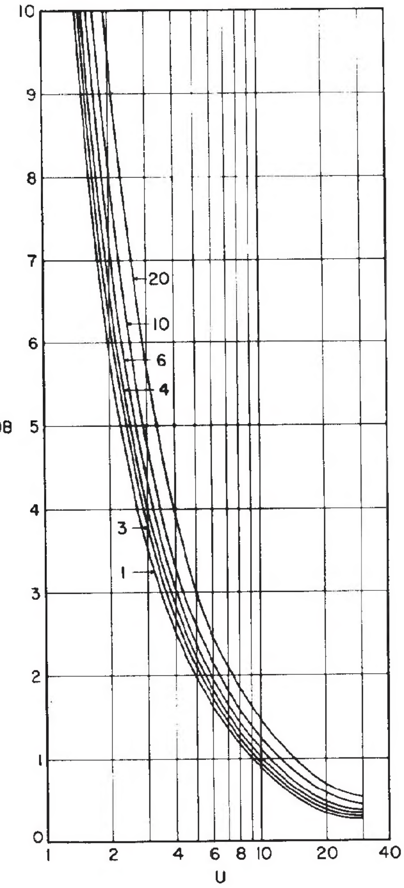

Fig. 2 Minimum insertion loss at midband in Butterworth filters of 1, 2, 3, 4, 6, 10 and 20 sections. The abscissa is the ratio x between the unloaded Q of each section and the minimum Q defined by Fig. 1.

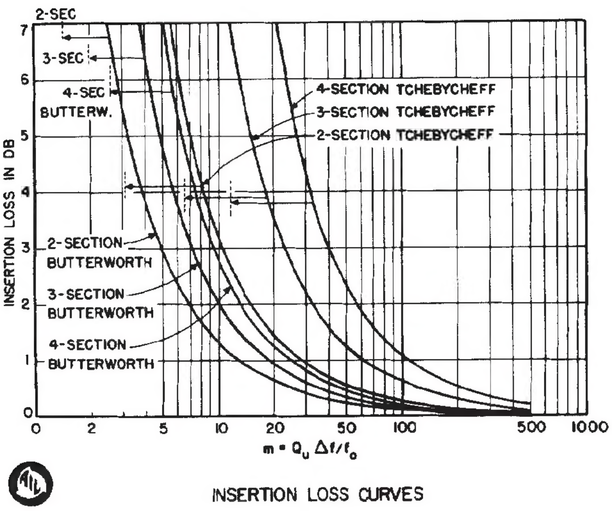

Fig. 3 Minimum midband insertion loss for a number of filters slotted as a function of the product m between the unloaded $Q = Q_u$ of each section times the relative bandwidth of the filter.

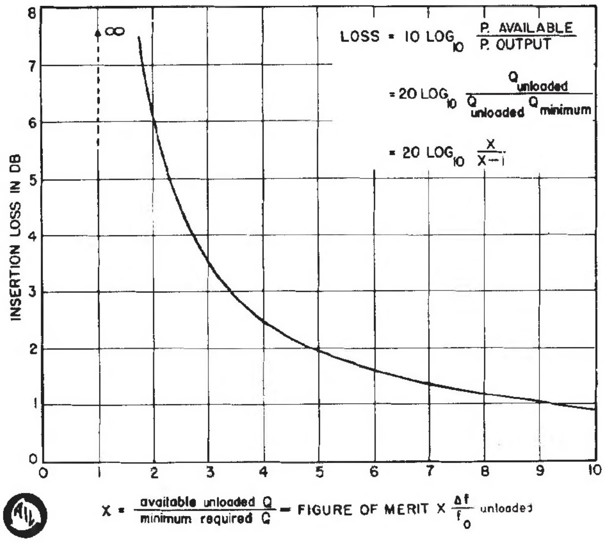

Fig. 5 General curve for minimum insertion loss. Minimum required Q can be found using Fig. 1.

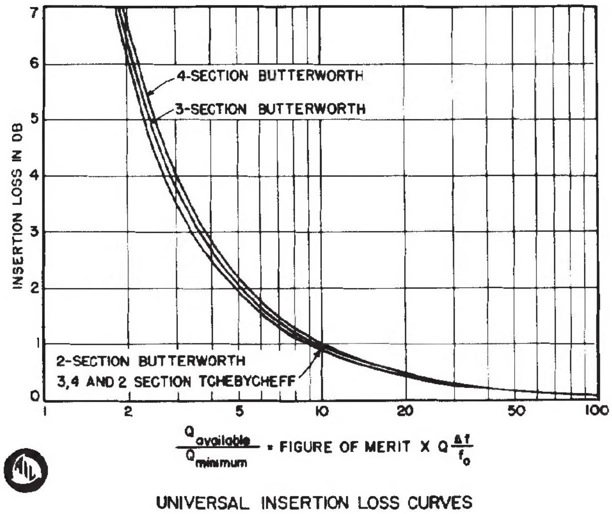

Fig. 4 The curves of Fig. 3 when normalized to the variable x = available unloaded Q/minimum unloaded Q.

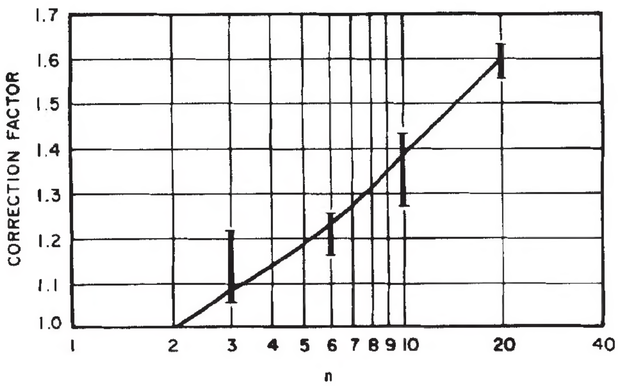

Fig. 6 Multiplying factor to be applied to the ordinates (db loss) of Fig. 5 to obtain a corrected value of loss if the number of sections is greater than two.

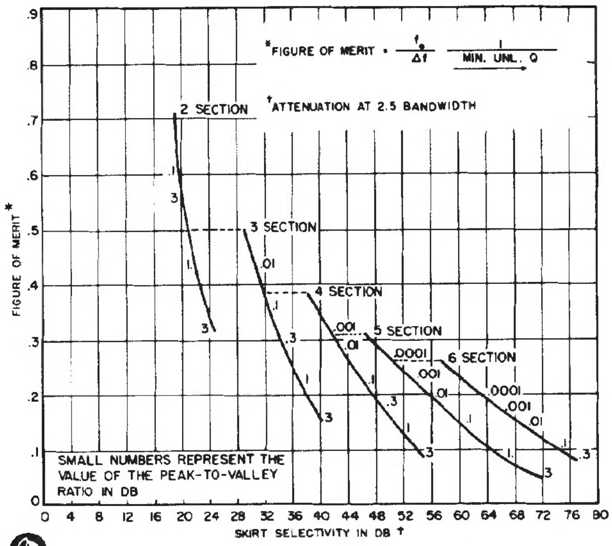

Fig. 7 The approximate quality of Tchebycheff and Butterworth filters as a function of skirt selectivity.

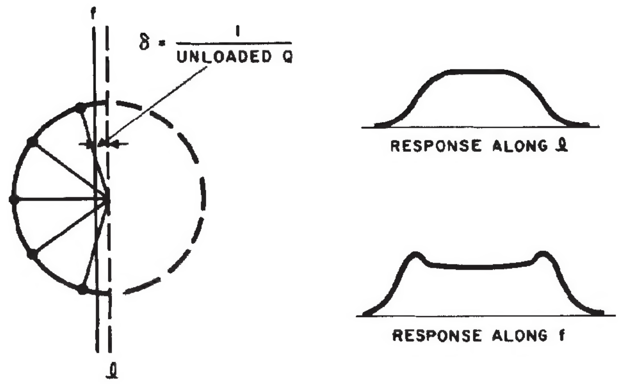

Fig. 8 The effect of finite losses in a Butterworth filter corresponds to shifting the circle of the poles toward the frequency axis by an amount δ equal to the reciprocal of the unloaded Q of each section.